

# An Experimental Comparison of Seriation Methods For One-Mode Two-Way Data

Michael Hahsler

*Southern Methodist University, Dallas, TX, USA*

---

## Abstract

Seriation aims at finding a linear order for a set of objects to reveal structural information which can be used for deriving data-driven decisions. It presents a difficult combinatorial optimization problem with its roots and applications in many fields including operations research. This paper focuses on a popular seriation problem which tries to find an order for a single set of objects that optimizes a given seriation criterion defined on one-mode two-way data, i.e., an object-by-object dissimilarity matrix. Over the years, members of different research communities have introduced many criteria and seriation methods for this problem. It is often not clear how different seriation criteria and methods relate to each other and which criterion or seriation method to use for a given application. These methods are represent tools for analytics and therefore are of theoretical and practical interest to the operations research community. The purpose of this paper is to provide a consistent overview of the most popular criteria and seriation methods and to present a comprehensive experimental study to compare their performance using artificial and a representative set of real-world datasets.

*Keywords:* Combinatorial optimization, heuristics, experimental study.

*2010 MSC:* 00-01, 99-00

---

## 1. Introduction

The field of analytics garners growing attention in the operations research community (Mortenson et al., 2015). A tool for analytics with roots in operations research is *seriation*, often also referred to as *sequencing* or *ordination*. Seriation arranges a set of objects into a linear order given available data with the goal to reveal structural information which can be used to support decision making. A typical assumption is that structural information can be revealed when similar objects are placed closer to each other than more dissimilar objects (Arabie & Hubert, 1996). To illustrate this idea, we apply seriation to the supreme court judges from the second Rehnquist U.S. Supreme Court (Sirovich, 2003) using their voting behavior. We would expect that the data contains a linear order from most conservative to most liberal judge. We measure pairwise dissimilarity by the

---

\*Cite this paper as: Michael Hahsler. An experimental comparison of seriation methods for one-mode two-way data. *European Journal of Operational Research*, 257:133–143, February 2017.

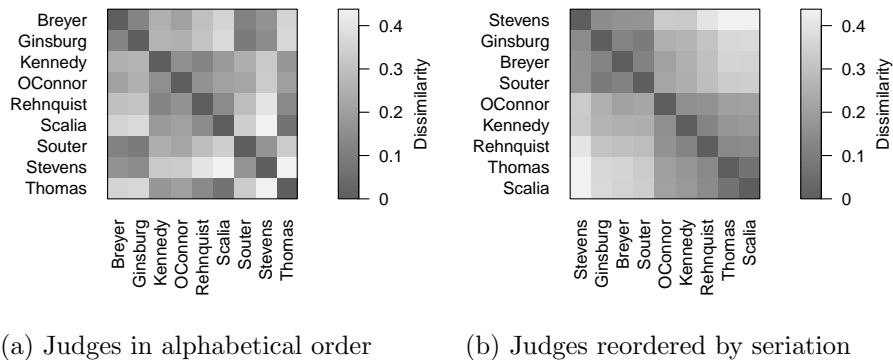


Figure 1: Dissimilarity matrix for Supreme Court judges in (a) original order and (b) after reordering using seriation.

joint probability of disagreement between the judges. The data is visualized in Figure 1 as a dissimilarity matrix with square color representing dissimilarity from low (dark) to high (light). Figure 1(a) shows the original data with the judges ordered alphabetically by last name, while Figure 1(b) has the judges rearranged using seriation. It is obvious that seriation succeeds in placing lower dissimilarity values (darker squares) closer to the diagonal ordering judges automatically from most conservative (Scalia) to most liberal (Stevens). The seriation also reveals more structural information by showing two darker blocks which represent two distinct groups, conservative and liberal judges, respectively.

Instead of judges, objects can for example be machines, products or customers. Seriation has a rich history starting at the turn of the 20th century with Petrie introducing the first formal method to find the chronological order for graves discovered in the Nile area given found objects (Petrie, 1899). The seriation problem was introduced to the operations research community as a combinatorial optimization task more than 40 years ago by McCormick and his colleagues when they studied different matrix reordering techniques. The group introduced the *bond energy algorithm (BEA)* (McCormick et al., 1972), a very influential method to identify natural groups in complex data matrices by simultaneously reordering columns and rows of a matrix such that an objective function called measure of effectiveness is maximized. This results in a matrix with groups of entries that are numerically as closely related to its four neighbors as possible. McCormick et al. (1972) present several applications ranging from finding relationships between marketing techniques and applications to factoring the problem of airport design into a number of smaller, more manageable subproblems. Lenstra (1974) showed that the optimization problem can be restated as two traveling-salesman problems, one for the rows and one for the columns. BEA and similar ideas are popular in the operations research literature for applications in manufacturing and especially for the important problem of optimal machine-part cell formation (see, e.g., Rogers & Kulkarni, 2005; Yang & Yang, 2008; Wu et al., 2010; Paydar & Saidi-Mehrabad, 2013; Boutsinas, 2013; Thanh et al., 2016). Seriation and the early work done by members of the operations research community is now also applied in such diverse fields as biology

(arrange gene expression data and read assembly), ecology (analyze plant associations), psychology (order subject-by-item response matrix), sociology (find group structure in sociograms), and visualization (reorder data tables, heatmaps and assess cluster tendency). Given the interdisciplinary nature of seriation and its applications, methods are developed and published by members of different research communities. A comprehensive recent historical overview of the development of seriation techniques and applications can be found in a review article by Liiv (2010).

This paper is based on our previous work (Hahsler et al., 2008) which discussed a smaller set of seriation criteria and methods mainly used for data visualization, and introduced the design of the open source R extension package seriation (Hahsler et al., 2016). Here we focus on a specific type of seriation problem, the problem of reordering dissimilarity matrices to reveal structural information. In the context of multidimensional scaling, this type of data is known as one-mode two-way data (Carroll & Arabie, 1980), indicating that the data only represents the relationship between a single set of objects using a two-dimensional object-by-object data array. Like in the case of BEA, seriation can also be performed directly on data matrices without first calculating dissimilarities as well as on two or higher-mode data where rows, columns and additional dimensions represent separate sets of objects which can be reordered simultaneously. Methods for two or higher-mode data are outside the scope of this paper and the interested reader is referred to the papers by Liiv (2010) and Hahsler et al. (2008).

The seriation methods considered in this paper are related to an iterative method called *moment ordering algorithm* developed by Deutsch & Martin (1971), two of McCormick’s colleagues, to identify a single dominant relationship in a data matrix that can be revealed by reordering the matrix. The authors also introduce a measure of effectiveness for the moment ordering algorithm, however it is not directly maximized by the algorithm. Similarly, the methods discussed in this paper try to find a good linear order, however, many directly optimize an objective function called a seriation criterion. Such optimization-based methods accept violations or deviations from a perfect linear order model in the data. This type of seriation is sometimes called statistical or probabilistic seriation to distinguish it from deterministic seriation often used in archaeological dating applications which

It is often not immediately clear how seriation criteria and methods proposed by authors from different research communities are related to each other and how different methods perform in terms of solution quality and runtime. The purpose of this paper is to (1) provide the operations research community with a review of the currently most popular seriation criteria and methods for one-mode two-way data using a consistent formulation as an combinatorial optimization problem, (2) organize seriation criteria and methods into groups, and (3) perform a rigorous experimental comparison study to highlight differences and provide some grounds for choosing the appropriate method for a given application. To enable operations research professionals to conduct experiments with their own data, we have implemented all methods discussed in this paper in the latest version of the R extension package seriation (Hahsler et al., 2016).

This paper is organized as follows. Section 2 formally introduces the seriation problem. Popular seriation criteria and seriation methods are reviewed in Sections 3 and 4, respectively. In Section 5 we present a comprehensive experimental study. We conclude the paper with Section 6.

## 2. The seriation problem

In this paper, we restrict the discussion to seriating or ordering a single set of  $n$  objects  $\mathcal{O} = \{O_1, \dots, O_n\}$  using as the input one-mode two-way data in the form of a  $n \times n$  symmetric dissimilarity matrix  $\mathbf{D} = \{d_{ij}\}$ , where  $d_{ij}$  for  $1 \leq i, j \leq n$  represents the pairwise dissimilarity between objects  $O_i$  and  $O_j$ , and  $d_{ii} = 0$  for all  $i$ . Similarities can be converted into dissimilarities using simple transformations, e.g.,  $d_{ij} = \frac{1}{1+s_{ij}}$ . Pairwise dis-(similarities) can be obtained in many ways. For example by calculating appropriate dissimilarity metrics (e.g., Euclidean distance), calculating correlation, estimating joint probabilities, or by obtaining pairwise similarity ratings from experts.

We define a permutation  $\pi : \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$  to indicate that objects originally ordered as  $O_1 O_2 \dots O_n$  will be reordered as  $O_{\pi(1)} O_{\pi(2)} \dots O_{\pi(n)}$ . A permutation function  $\psi_\pi : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$ , which reorders  $\mathbf{D}$  according to permutation  $\pi$ , can be written as  $\psi_\pi(\mathbf{D}) = \{d_{\pi(i), \pi(j)}\} = \mathbf{P}_\pi \mathbf{D} \mathbf{P}_\pi^T$ , where  $\mathbf{P}_\pi$  is the permutation matrix corresponding to  $\pi$ . The permutation matrix  $\mathbf{P}_\pi$  is defined as the identity matrix  $\mathbf{I}_n$  with rows and columns reordered according to the permutation  $\pi$ , i.e.,  $\mathbf{P}_\pi = \psi_\pi(\mathbf{I}_n)$ . We use  $\Psi$  to denote the set of all possible permutation functions.

The quality of the arrangement of objects is assessed by a given criterion. Without loss of generality, we use a loss function  $L$  here (a merit function  $M$  can be easily transformed into a loss function). This leads to the following optimization problem.

$$\begin{aligned} & \text{minimize} && Z = L(\psi_\pi(\mathbf{D})) \\ & \text{subject to} && \psi_\pi \in \Psi \end{aligned}$$

Finding the optimal permutation is in general a hard discrete optimization problem with a set of feasible solutions  $\Psi$  of size  $O(n!)$ . We will discuss complexity in more detail when we introduce seriation methods in Section 4.

Seriation is related to *unidimensional scaling* with equal weights. Unidimensional scaling (Mair & De Leeuw, 2015) is the one-dimensional special case of multidimensional scaling with the objective to find for each object a coordinate along a line while minimizing stress given by the difference between the original pairwise dissimilarities between objects and the distance of their coordinates along the line. Interestingly, the methods used for unidimensional scaling are very different from the methods applied for the general case of multidimensional scaling (Mair & De Leeuw, 2015) leading to a combinatorial problem similar to seriation. However, seriation is only concerned with finding the order of objects along the line, but not the actual coordinates.

Seriation is also related to *rankings without ties* which impose a strict total order on the set of objects (Davey & Priestley, 1990). However, while the direction of the order (i.e., first place, second place, etc.) is important for an ordered set  $\langle \Omega; \langle \rangle$ , for seriation, an order and its exact reverse are equivalent. That is,  $O_1 < O_2 < \dots < O_n \equiv O_n < O_{n-1} < \dots < O_1$ . For the example with the Supreme Court judges in Figure 1 this means that the found order from Scalia to Stevens and the exact reverse from Stevens to Scalia are equivalent. This is especially important when comparing different seriation results. Orders are often compared using rank-order correlation coefficients like Kendall's tau (Kendall, 1938), which evaluate for each pair of objects, if they are in the same order in both rankings. For comparing two orders, which are exactly the same, the correlation coefficient will be +1. For unrelated orders the correlation is close to zero. However, for

an order and its exact reverse, the coefficient is  $-1$ . It is easy to see, that the similarity between two seriation orders can be measured using the absolute value of the correlation coefficient. Recently, Goulermas et al. (2016) have introduced a specialized measure to compare seriation orders, called positional proximity measure. This measure compares the distance of each pair of objects in the two orders and thus is not affected by a reversal of the order.

In the following, we will introduce some popular seriation criteria and seriation methods.

### 3. A review of seriation criteria

Finding a good seriation order, where similar objects are close to each other and dissimilar objects are more distant, is equivalent to finding a dissimilarity matrix where small dissimilarity values are arranged close to the main diagonal and large values are pushed far away. There are several ways to construct a criterion formalizing this idea. We organize the most popular criteria in this paper by the way they are constructed into groups based on gradient conditions, agreement between object rank differences and dissimilarities, and path length. Table 1 summarizes the definitions of the discussed criteria calculated for the current order of dissimilarity matrix  $\mathbf{D} = \{d_{ij}\}$ . The used indicator and sign functions are defined as  $I(x > y) = 1$  if  $x > y$  and 0 otherwise; and  $\text{sign}(x) = +1$  if  $x > 0$ , 0 if  $x = 0$  and  $-1$  if  $x < 0$ , respectively.

#### 3.1. Gradient conditions

The perfectly ordered dissimilarity matrix is called an *anti-Robinson matrix* after the statistician W.S. Robinson (1951). Here the dissimilarity values in all rows and columns monotonically increase when moving away from the main diagonal, indicating that more similar objects are always placed closer together. For most real data it is very unlikely that a permutation function exists which will result in a perfect anti-Robinson matrix. Hubert et al. (2001) formalized the idea of measuring the closeness of a matrix to the anti-Robinson form by defining gradient conditions

$$\begin{aligned} \text{within rows} \quad & d_{ik} \leq d_{ij} \quad \text{for} \quad 1 \leq i < k < j \leq n \quad \text{and} \\ \text{within columns} \quad & d_{kj} \leq d_{ij} \quad \text{for} \quad 1 \leq i < k < j \leq n. \end{aligned}$$

Row and/or column gradient conditions are the basis of several seriation criteria. Chen (2002) counts the number of violations of the gradient conditions. He called these violations *anti-Robinson events (AR events)*. AR events can also be weighted by the magnitude of the violation called *anti-Robinson deviations*. We can also count agreements in addition to violations. The difference between agreements and violations is called the *gradient measure* (Hubert et al., 2001). As for AR events, the gradient measure can also be weighted by the magnitude of agreements and violations resulting in a *weighted gradient measure*.

The seriation measures discussed so far are concerned with revealing global structure in the data by optimizing over the whole matrix. That is, for each object the distance to all other objects is considered. For some applications, it can be useful to reveal localized structures by only considering the neighborhood of each object. Such a criterion is

| Measure  | Definition  |
|--|---|
| <b>Gradient conditions</b>   |   |
| Anti-Robinson (AR) events (Chen, 2002)                               | $\sum_{i < k < j} f(d_{ik}, d_{ij}) + f(d_{kj}, d_{ij})$ ,<br>with $f(x, y) = I(x > y)$   |
| AR deviations (Chen, 2002)   | with $f(x, y) =  y - x  I(x > y)$   |
| Gradient measure (Hubert et al., 2001)                               | with $f(x, y) = -\text{sign}(y - x)$  |
| Weighted gradient measure (Hubert et al., 2001)                      | with $f(x, y) = - y - x  \text{sign}(y - x)$  |
| Relative generalized Anti-Robinson events (RGAR) (Tien et al., 2008) | $\frac{1}{m} \sum_{i=1}^n \left( \sum_{(i-w) \leq i < k < j} I(d_{ik} < d_{ij}) \right.$<br>$\left. + \sum_{i < k < j \leq (i+w)} I(d_{kj} > d_{ij}) \right)$ ,<br>with window size $1 < w < n$<br>and $m = (\frac{2}{3} - n)w + nw^2 - \frac{2}{3}w^3$ |
| <b>Rank/dissimilarity agreement</b>                                  |   |
| Least squares criterion (Caraux & Pinloche, 2005)                    | $\sum_{i,j=1}^n (d_{ij} -  i - j )^2$   |
| Inertia criterion (Caraux & Pinloche, 2005)                          | $-1 \times \sum_{i,j=1}^n d_{ij} (i - j)^2$   |
| 2-Sum criterion (Barnard et al., 1993)                               | $\sum_{i,j=1}^n \frac{1}{1+d_{ij}} (i - j)^2$   |
| Linear seriation criterion (LS) (Hubert & Schultz, 1976)             | $-1 \times \sum_{i,j=1}^n d_{ij}  i - j $   |
| Banded anti-Robinson form (BAR) (Earle & Hurley, 2015)               | $\sum_{ i-j  \leq b} d_{ij} (b + 1 -  i - j )$<br>with band width $1 \leq b < n$  |
| <b>Path length</b>   |   |
| Hamiltonian path length (PL) (Hubert, 1974; Caraux & Pinloche, 2005) | $\sum_{i=1}^{n-1} d_{i,i+1}$  |

Table 1: Popular seriation criteria.

165 *relative Generalized Anti-Robinson events (RGAR)* (Tien et al., 2008) which only counts  
 AR events in a band (a window specified by  $w$ ) around the main diagonal of the reordered  
 dissimilarity matrix and normalizes the sum by the maximum number of possible events  
 in the band. RGAR can be used to create a tradeoff between only looking at local  
 structure (neighboring objects) with  $w = 2$  and global structure with  $w = n - 1$  (which  
 170 is a scaled equivalent to anti-Robinson events as defined above).

### 3.2. Agreement between object rank differences and dissimilarities

A good seriation can also be described as an order where the dissimilarities between  
 objects agree with their rank difference in the order. That is, objects placed farther apart  
 also are more dissimilar. Several criteria can be created to evaluate this agreement. The  
 175 *least squares criterion* (Caraux & Pinloche, 2005) uses the squared difference between all  
 pairwise dissimilarities and the rank differences. It is similar to the objective function  
 of unidimensional scaling with equal weights where the distance between coordinates is  
 used instead of the rank difference (Mair & De Leeuw, 2015). *Inertia* (Caraux & Pin-  
 loche, 2005) focuses on how far large dissimilarity values are pushed away from the main  
 180 diagonal. On the contrary, the *2-Sum criterion* (Barnard et al., 1993) penalizes pushing  
 high similarities away from the diagonal. The *linear seriation criterion (LS)* (Hubert &  
 Schultz, 1976) is related to the inertia criterion, but does not square the rank differences,  
 and therefore does not emphasize the impact of the distances between objects that are  
 placed far from each other. Earle & Hurley (2015) introduced a criterion equivalent to  
 185 LS (scaled by  $1/2$ ) and call it anti-Robinson criterion (ARc). Earle & Hurley (2015) also  
 introduced a banded version of ARc which only considers the agreement between the  
 rank difference and the dissimilarities in a band of width  $1 \leq b < n$  around the main  
 diagonal, and thus follows the same idea of revealing localized structure as RGAR. For  
 $b = 1$ , the criterion is equivalent to the Hamiltonian path length (see below) and with  
 190  $b = n - 1$  it is equivalent to ARc/LS.

### 3.3. Path length

A dissimilarity matrix can be viewed as a finite weighted complete graph  $G = (V, E)$ ,  
 where vertices are the set of objects, i.e.,  $V(G) = \{O_1, O_2, \dots, O_n\}$  and each edge  $e_{ij} \in$   
 $E(G)$  is labeled with a weight given by the dissimilarity  $d_{ij}$ . Finding a linear order  
 195 can be seen as a *Hamiltonian path* that visits each object exactly once. Minimizing  
 the Hamiltonian path length results in a seriation optimal with respect to the local  
 structure given only by dissimilarities between neighboring objects (Hubert, 1974; Caraux  
 & Pinloche, 2005).

### 3.4. Computational complexity

200 It is easy to see from the definitions that each group of measures has a different compu-  
 tation complexity. Gradient conditions take  $O(n^3)$  to compute, while rank/dissimilarity  
 agreement takes  $O(n^2)$ , and path length can be computed in  $O(n)$ .

## 4. A review of seriation methods

205 Seriation is a discrete optimization problem which, in the most general case, involves  
 evaluating all feasible solutions. Due to the combinatorial nature, the number of possible

| Technique  | Objective function                                  |
|--|---|
| <b>Criterion optimization methods</b>  |   |
| Integer linear programming (ILP) (Brusco, 2002)  | Gradient conditions                                 |
| Dynamic programming (Hubert et al., 2001)  | Gradient conditions                                 |
| Branch-and-bound (Brusco & Stahl, 2005)  | Gradient conditions                                 |
| Genetic algorithm (Goldberg, 1989; Soltysiak & Jaskulski, 1998)  | Various   |
| Simulated annealing (ARSA) (Brusco et al., 2008)   | Linear seriation (mistake in the published version) |
| Spectral seriation (Atkins et al., 1999; Ding & He, 2004; Fogel et al., 2014)  | 2-Sum criterion                                     |
| TSP solver (various) (Wilkinson, 1971)   | Hamiltonian path length                             |
| Quadratic assignment problem heuristic (QAP) (Hubert & Schultz, 1976; Caraux & Pinloche, 2005; Goulermas et al., 2016) | 2-Sum criterion, linear seriation, inertia or BAR   |
| <b>Dendrogram methods</b>  |   |
| Hierarchical clustering (HC) (Eisen et al., 1998)  | Other (depends on linkage)                          |
| Gruvaeus and Wainer reordering (GW) (Gruvaeus & Wainer, 1972)  | Restricted path length                              |
| Optimal leaf ordering reordering (OLO) (Bar-Joseph et al., 2001)   | Restricted path length                              |
| DendSer reordering (Earle & Hurley, 2015)  | Various (restricted)                                |
| <b>Other methods</b>   |   |
| Multidimensional scaling (MDS) (Kendall, 1971)   | Other (stress)                                      |
| Rank-two ellipse seriation (R2E) (Chen, 2002)  | None  |
| Sorting Points Into Neighborhoods (SPIN) (Tsafrir et al., 2005)  | Other (energy)                                      |
| Visual Assessment of Tendency (VAT) (Bezdek & Hathaway, 2002)  | Other (MST)   |

Table 2: Popular seriation techniques.

solutions grows with problem size (number of objects,  $n$ ) by the order  $O(n!)$ . Seriation has been shown to be an NP-complete problem (George & Pothen, 1997), and many heuristic methods have been proposed. We organize methods here into methods that directly try to optimize a seriation criterion, dendrogram-based methods, and other methods which produce good seriations without directly targeting a specific seriation criterion. Table 2 summarizes popular methods and indicates for each what, if any, seriation criterion is optimized.

#### 4.1. Seriation criterion optimization methods

The large, discrete search space makes a brute-force enumerative approach infeasible for all but very small problems. Seriation problems using linear loss functions can be formulated as integer linear programs (ILPs) and solved using standard ILP solvers. Brusco (2002) discusses ILP formulations for the number of anti-Robinson events and



for gradient measures and concludes that they are useful for very small problems. To solve somewhat larger problems, partial enumeration methods can be used. For example, dynamic programming (Hubert et al., 2001) and branch-and-bound strategies (Brusco & Stahl, 2005) were used to optimize the unweighted and weighted gradient measure. However, these methods are still limited to very moderate sizes of up to 40 objects (Hahsler et al., 2008). The following methods have been proposed for larger problems.

Metaheuristics like *genetic algorithms* (Soltysiak & Jaskulski, 1998) and *simulated annealing* (Brusco et al., 2008) have been used. For genetic algorithms, many genetic operators proposed for the traveling salesperson problem can also be used for seriation problems. Examples are ordered crossover and simple reversal and swap mutation operators (Goldberg, 1989).

*Spectral seriation* uses a relaxation to minimize the 2-Sum criterion (Barnard et al., 1993). Rewriting the minimization problem using a permutation vector, its rescaled inverse, and a Lagrangian multiplier for the constraint allows us to recover the optimal order from the Fielder vector, i.e., the second smallest eigenvector of the Laplacian of the similarity matrix.

To minimize the Hamiltonian path length is related to the *traveling salesman problem (TSP)*, which is a well known and well researched combinatorial optimization problem with a large set of heuristics and exact methods (see, e.g., Gutin & Punnen, 2002). The TSP results in a circular tour, but the problem can be easily transformed into a linear order problem, by inserting a dummy object which is infinitely distant from all other objects (Garfinkel, 1985). Cutting the tour at the dummy object results in the desired path.

Hubert & Schultz (1976) showed that optimizing the linear seriation criterion can be rewritten as a type of facility location problem called the *Quadratic Assignment Problem (QAP)*

$$\text{QAP}(\mathbf{A}, \mathbf{B}) : \min_{\pi} \sum_{i,j=1}^n a_{ij} b_{\pi(i),\pi(j)},$$

where the objective is to find the optimal assignment for  $n$  facilities to exactly one of  $n$  locations each. Flows between the facilities are given by flow matrix  $\mathbf{A}$  and the relative position of the locations is represented by distance matrix  $\mathbf{B}$ . The objective is to minimize transportation cost given by the sum of all flows times the corresponding distances. By defining a flow matrix depending on the relative position of objects in the seriation order, optimizing the linear seriation criterion can be reformulated as a QAP, i.e.,

$$\min_{\pi} \sum_{i,j=1}^n d_{\pi(i),\pi(j)} - |i - j| \quad \text{leads to} \quad \text{QAP}(\{-|i - j|\}_{n \times n}, \mathbf{D}).$$

As seriation itself, the QAP is in general also NP-hard, but methods including QIP, linearization, branch and bound and cutting planes as well as heuristics including Tabu search, simulated annealing, genetic algorithms, and ant systems can be used to find good solutions (Burkard et al., 1999). Barnard et al. (1993) formulate the 2-Sum problem as  $\text{QAP}(\{(i - j)^2\}_{n \times n}, \mathbf{S})$ , where the similarity matrix is defined as  $\mathbf{S} = \frac{1}{1+\mathbf{D}}$ . Similarly, it is easy to see that optimizing inertia and the BAR criterion can also be formulated

as QAPs. Inertia leads to QAP( $\{-(i-j)^2\}_{n \times n}, \mathbf{D}$ ). For BAR we can define the flow matrix

$$\mathbf{A}_{\text{BAR}} = \begin{cases} b + 1 - |i - j| & \text{if } |i - j| \leq b, \\ 0 & \text{otherwise.} \end{cases}$$

Then optimizing BAR can be formulated as QAP( $\mathbf{A}_{\text{BAR}}, \mathbf{D}$ ).

#### 4.2. Dendrogram methods

*Hierarchical clustering* produces a series of nested clusterings which can be visualized by a dendrogram. A dendrogram is a binary tree where the leaf nodes represent the individual objects and each internal node represents joining objects into larger groups of similar objects till all objects are joined in the tree's root. For an example, see Figure 2(a) later in this paper. As a simple and fast heuristic to find a linear order of objects, the order of the leaf nodes in a dendrogram structure can be used (Eisen et al., 1998). This method does not directly optimize a seriation criterion, however, it can be used as a starting point. Subtrees can be rotated without changing the nested cluster structure, and the original leaf node order is typically an artifact of the used clustering algorithm. To improve the presentation of the dendrogram, several methods for rotating subtrees to minimize an objective function under the constraints given by the dendrogram have been proposed. Gruvaeus & Wainer (1972) suggest to obtain a unique order by requiring to order the leaf nodes such that at each level the objects at the edge of each cluster are adjacent to that object outside the cluster to which it is nearest and they provide a simple heuristic. Bar-Joseph et al. (2001) developed an efficient procedure to rearrange the dendrogram such that the Hamiltonian path connecting the leaves is minimized and called this the *optimal leaf order*. Earle & Hurley (2015) recently developed a general framework for dendrogram seriation which is able to use various criteria (e.g., path length, banded anti-Robinson form, linear seriation criterion). The algorithm applies node operators (subtree translation and/or rotation) in a greedy fashion to quickly find solutions of appropriate quality for visualization purposes.

#### 4.3. Other methods

*Multidimensional Scaling (MDS)* tries to find a lower-dimensional representation of the similarity structure between objects by creating new, so-called principle coordinates, while minimizing the stress (i.e., the squared difference between the dissimilarity of two objects in the lower-dimensional and the original space). Although seriation is more closely related to unidimensional scaling, MDS can be used to find reasonable orderings (Kendall, 1971). Objects can be ordered along the first principal coordinate obtained from metric or non-metric MDS. Alternatively, the objects can be projected onto the first two principal coordinates found by MDS and then ordered by the angle in this space. The order is then split by the largest angle gap between adjacent objects (Friendly, 2002).

Chen (2002) proposes the *rank-two ellipse seriation* procedure. For this method, a sequence of correlation matrices starting with the dissimilarity matrix is generated. Once the rank of a generated correlation matrix drops to two, all objects are projected onto the two eigenvectors of this matrix. The objects form an ellipse which can be used to extract a seriation order.

295 *Visual Assessment of Tendency (VAT)* (Bezdek & Hathaway, 2002) was developed  
as a visual method to judge if a dataset should be clustered. It creates an order based on  
Prim’s algorithm for finding a minimum spanning tree (MST) in a weighted connected  
graph representing the dissimilarity matrix. The order in which the objects are added to  
the MST represents a seriation order. This method is related to single-link hierarchical  
300 clustering.

*Sorting Points Into Neighborhoods (SPIN)* (Tsafrir et al., 2005) tries to minimize the  
energy for the permutation matrix using a weight matrix which depends on the rank  
difference of objects. The authors suggest two algorithms, the Side-to-Side algorithm  
(STS) which tries to push out large dissimilarity values, and the neighborhood algorithm  
305 (NH) which concentrates low dissimilarity values around the diagonal.

## 5. Experimental comparison

The goal of this experimental comparison is to produce results which are generalizable  
to a wide variety of application areas. To achieve this goal, we have chosen to use  
two artificial datasets and ten real-world datasets. After introducing the datasets, we  
310 will use them to compare popular seriation criteria and the results of different seriation  
techniques. We will conclude with identifying the most efficient methods and present  
scalability results.

### 5.1. Datasets

We will use two types of simulated datasets.

- 315 • **Pre-Robinson data:** We create randomly permuted perfect anti-Robinson dis-  
tance matrices by reversing the process of unidimensional scaling. We randomly  
pick  $n$  coordinates on a line and then calculate pairwise distances to create the  
distance matrix. Since these matrices contain a perfect linear order, they represent  
easy seriation problems Laurent & Seminaroti (2015).
- 320 • **Random data:** These matrices are created as distance matrices between sets of  
 $n$  objects randomly placed into two-dimensional Euclidean space. They represent  
hard seriation problems with no apparent linear structure present.

Table 3 summarizes the used real-world datasets. Most datasets are also available in  
the R extension package *seriation* (Hahsler et al., 2016). The datasets come from very  
325 different areas including archaeology, psychology, political science, biology and social  
media. The dataset size ranges from 24 to 229 objects. We will later perform scalability  
experiments using random data with up to 10,000 objects.

### 5.2. Relationship between seriation criteria

Many seriation criteria have been proposed and we have organized them in this paper  
330 into gradient condition, rank/dissimilarity agreement and path length-based. For the  
criteria which have a parameter (RGAR and BAR), we use for the window/band width  
the minimum value, 20% of the number of objects (which was suggested as the default  
value for BAR (Earle & Hurley, 2015)) and the maximum value.

| Name        | Description (Source)   | $n$ |
|-------------|--|-----|
| Psych24     | Pearson correlation between results of 24 psychological tests given to 145 seventh and eighth grade students in a Chicago suburb (Holzinger & Swineford, 1939).                | 24  |
| Irish       | Euclidean distances of scaled results of eight referenda for 41 Irish communities (de Falguerolles et al., 1997).  | 41  |
| Munsingen   | Jaccard index for incidence matrix for 59 graves and 70 artifacts (Hodson, 1968).  | 59  |
| Votes       | Jaccard index for 16 key votes for a sample of 100 of 435 congress men (1984) encoded as 32 binary features (Lichman, 2013).   | 100 |
| Zoo         | Euclidean distance for 17 features for 101 animals (Lichman, 2013).  | 101 |
| Iris        | Euclidean distances (scaled) for Fisher’s Iris dataset with 150 flowers and four features (Fisher, 1936).  | 150 |
| Wood        | Euclidean distance for sample of the normalized gene expression data for six locations in the stem of Popla trees (Hertzberg et al., 2001).                                    | 136 |
| Elutriation | Ratios of gene expression levels for a sample of genes of <i>Saccharomyces cerevisiae</i> with 14 eigengenes (from SVD) as features. (Euclidean distance) (Alter et al., 2000) | 200 |
| Facebook    | Sample of individuals from an ego-network. (count of connections of length 1 and 2) (Leskovec & Krevl, 2014)   | 200 |
| DBLP        | Members of the first 10 communities in the computer science bibliography co-authorship network. (count of connections of length 1 and 2) (Leskovec & Krevl, 2014)              | 229 |

Table 3: Real-world datasets.

To experimentally establish the relationship between seriation criteria, we create 100  
 335 random dissimilarity matrices for 100 objects each, and calculate the value of each crite-  
 rion. Then we rank the 100 dissimilarity matrices according to each criterion. Seriation  
 criteria are similar if they rank the matrices in a similar way, i.e., agree on which matrix  
 is closer to a good seriation order. To measure pairwise similarity between seriation crite-  
 ria, we calculate Kendall’s tau rank correlation coefficient from the way each criterion  
 340 ranks the 100 matrices resulting in a criterion-to-criterion correlation matrix. To reveal  
 structural information in this matrix, we will use a dendrogram-based seriation tech-  
 nique on a distance matrix obtained by subtracting the correlations from one. We use  
 hierarchical clustering with Ward’s minimum variance method which is known to lead  
 to compact clusters (Ward, 1963) and then apply optimal leaf ordering. The resulting  
 345 dendrogram is shown in Figure 2(a). Two seriation criteria are more similar, if they are  
 joined in the dendrogram at a lower height. We also show the seriated correlation matrix  
 in Figure 2(b). Note that we only show positive correlations, since we are interested in  
 the similarity between criteria.

Interestingly, the experiments do not just show the three groups of seriation criteria  
 350 described in Table 1, but a more complicated structure. In the dendrogram in Figure 2(a),  
 the criteria fall into four groups.

- **Group 1:** Path length is equivalent to BAR with the minimal band width of  $b = 2$   
 and also similar to RGAR with minimal and 20% window size.
- **Group 2:** BAR with the maximal window size is equivalent to the linear seri-  
 355 ation (LS) criterion. And both are similar to BAR width a band with of 20%.  
 Interestingly, all these rank/dissimilarity agreement criteria are very similar to AR  
 deviations which is a weighted measure constructed using gradient conditions.
- **Group 3:** This group contains all gradient condition criteria except AR deviations  
 and the RGAR with the maximal window sizes.
- **Group 4:** Contains the rank/dissimilarity agreement measures 2-Sum, Inertia and  
 360 least squares.

The four groups are the result of seriation and thus the order is also meaning full.  
 It represent an order from focusing on local structure (group 1) to emphasizing global  
 structure (group 4). The strong emphasis in group 4 results from the fact that the  
 365 rank differences are squared in all criteria in this group. Groups 2 and 3 represent an  
 intermediate step that considers global structure, but do not overemphasize the influence  
 of objects that are placed very far from each other.

The seriated correlation matrix in Figure 2(b) shows the same information, however,  
 it presents a clearer picture showing that AR deviations is correlated with groups 2 and  
 370 3. Clustering has placed it into group 2, but by construction it should be part of group 3.  
 It also indicates that although the dendrogram splits groups 2 and 3, they are actually  
 related with each other by forming a clearly visible block.

### 5.3. Comparison of seriation methods

Next, we experimentally compare the solution quality produced by different seriation  
 375 methods. The procedure is to apply each seriation method to a set of datasets, and

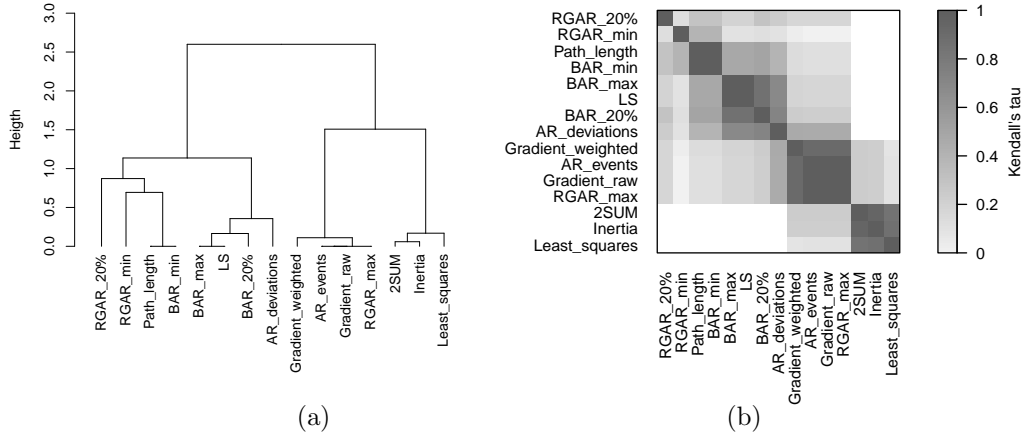


Figure 2: Similarity between seriation criteria as a (a) dendrogram and (b) an seriated correlation matrix.

then compute the value of different criteria for the found solutions. This is especially interesting for methods which do not explicitly optimize a seriation criterion like the dendrogram methods without reordering, MDS, rank-two ellipse seriation, SPIN and VAT.

380 In this paper we report the results for two prototypical seriation criteria, namely, anti-Robinson events and path length. Experiments not reported in this paper show that the results within each of the path length-based group and the group formed by gradient-based and rank/dissimilarity agreement methods are very consistent. We compare the solutions for the methods shown in Table 2. We exclude ILP, dynamic programming and  
 385 branch-and-bound because of their extremely limited scalability. Genetic algorithms are also excluded because of the excessive runtime for the used datasets. We implemented the exact algorithms described in the given references. As a fast TSP heuristic we use the best solution found in ten runs of the arbitrary insertion heuristic followed by complete 2-opt local search (Hahsler & Hornik, 2016). Experimentation showed that  
 390 this combination produces generally good seriation results. To solve the QAP we use the simulated annealing heuristic described by Burkard & Rendl (1984). Implementation details can be found in Hahsler et al. (2016).

We start with pre-anti-Robinson matrices which contain a perfect linear order. The result for 10 random pre-anti-Robinson matrices with 100 objects each is shown in the  
 395 box plot in Figure 3(a). The box plot shows for each method the criterion's median value (horizontal bar inside the box), the interquartile range (box) and outliers (circles) for the 10 runs. We use plots since the experiments produce too many values to present them in table form. Methods are sorted from best to worse median value and ties are broken alphabetically. We see that many methods perfectly recover the linear structure in the  
 400 data resulting in no anti-Robinson events or the minimal possible path length. This is expected because pre-anti-Robinson matrices contain a perfect linear order and seriating them is known to be an easy problem (Laurent & Seminaroti, 2015).

A much more difficult problem is to seriate random data where no linear order is present. We create 10 dissimilarity matrices from 100 objects randomly placed in two-

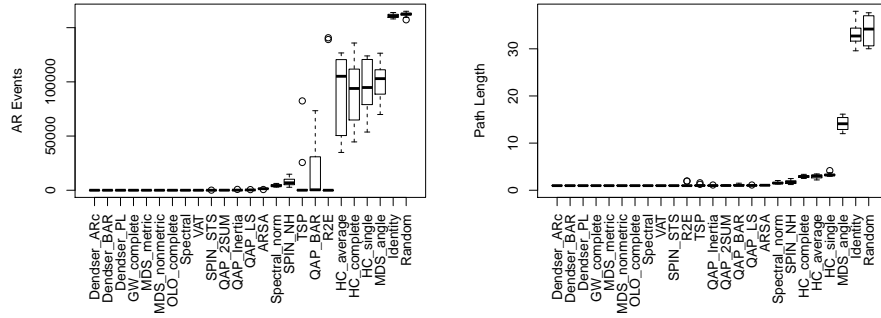
405 dimensional Euclidean space. The results are shown in Figure 3(b). We report for path  
length the relative optimality gap where the optimal path length was obtained using  
the Concorde TSP solver (Applegate et al., 2006). Finding the optimal solution for AR  
events is not feasible and we report the gap to the best found solution instead. In terms  
of anti-Robinson events, simulated annealing (ARSA), MDS, QAP and spectral seriation  
410 perform very well while hierarchical clustering and TSP-based methods perform poorly.  
For path length it is exactly the opposite with TSP and hierarchical clustering with  
reordering to reduce path length performing the best. These results are expected since  
each of the two groups optimizes for one of the two groups of criteria. The only exception  
is the good performance of MDS which minimizes stress rather than a criterion related  
415 to anti-Robinson events or the gradient criterion. While Kendall (1971) argued that  
MDS can be used to find reasonable seriation orders, others report that using MDS to  
find a single dimension is prone to ending up in local optima (Mair & De Leeuw, 2015).  
However, in our experiments MDS performed similar to the best other seriation methods.

To compare the performance on real-world data, we use the ten real-world datasets  
420 introduced above. We report again the relative optimality gap for path length and the  
gap to the best found solution for AR events. The results in Figure 3(c) are very similar  
to the results obtained for random data which indicates that real-world data represent  
difficult seriation problems. Since the random and identity orders have such a large gap,  
the boxes are cut off in the figure.

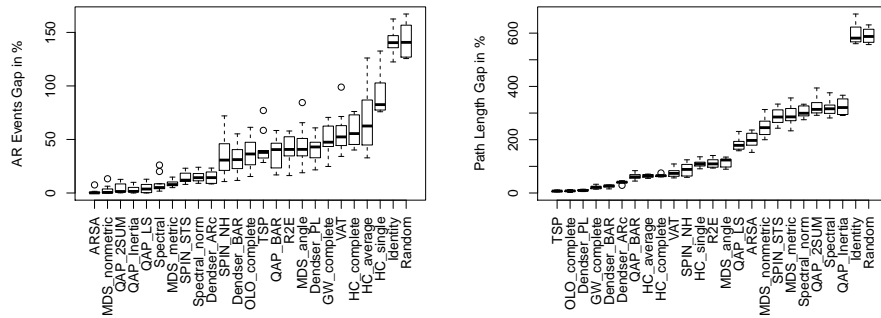
425 Next, we investigate how similar the resulting seriation orders produced by different  
methods are. We apply all seriation methods to a dissimilarity matrix for random data  
(100 objects) and then compare the resulting orderings pairwise using the absolute value  
of Kendall’s rank order coefficient. We use the absolute value since a seriation order  
and its reverse are equivalent. We repeat this with 10 random datasets and average  
430 the pairwise correlations. For visualization we use again a dendrogram obtained using  
hierarchical clustering with Ward’s minimum variance criterion and optimal leaf ordering.  
Figure 4(a) shows the dendrogram. The methods fall into tree groups.

- **Group 1:** This group contains all methods based on complete-link hierarchical  
435 clustering with and without different reordering strategies (DendSer methods, OLO  
and GW).
- **Group 2:** Contains simulated annealing (ARSA), most QAP-based methods (LS,  
2-Sum and inertia), metric/nonmetric MDS, spectral seriation and the neighbor-  
hood SPIN algorithm.
- **Group 3:** Consists of the remaining methods which are only very loosely related  
440 to each other and the methods in group 2. As expected, identity and random order  
are not related to any other method.

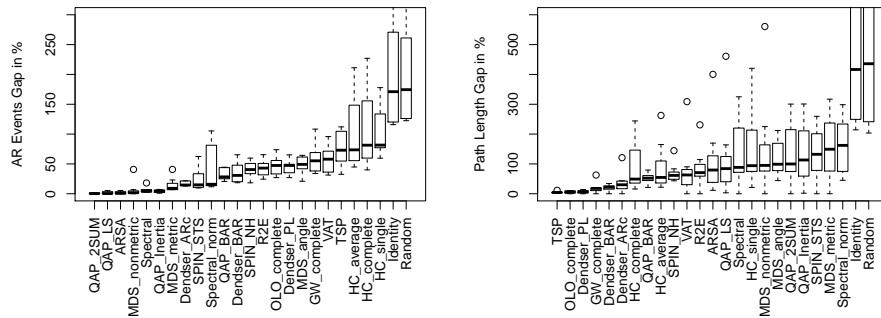
The seriated correlation matrix in Figure 4(b) shows the same result with two darker  
blocks forming groups 1 and 2. These results are interesting since they mean that there  
are groups of algorithms which produce relatively similar seriation results and, if runtime  
445 or scalability are important for the application then we can use the fastest, most scalable  
algorithm in the respective group.



(a) Pre-Robinson data.



(b) Gap in % on random data.



(c) Gap in % on the ten real datasets.

Figure 3: Anti-Robinson events and Hamiltonian path length for different seriation methods on (a) a Pre-Robinson matrix, (b) random data, and (c) ten real datasets.



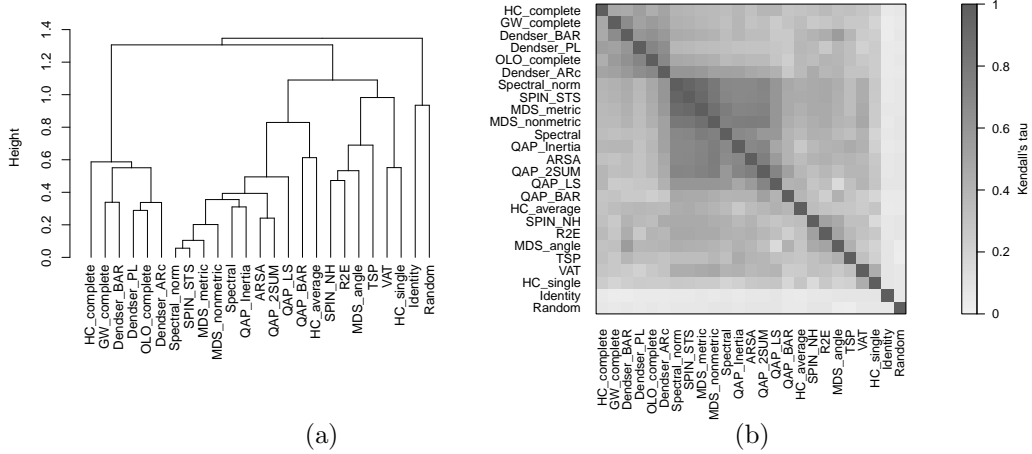


Figure 4: Similarity between orders produced by different seriation methods as (a) a dendrogram and as a (b) seriated correlation matrix.

#### 5.4. Runtime, efficiency and scalability

Runtime and scalability are very important for some applications. For example, for real-time visualization, reordering needs to be performed almost instantaneously and for arranging gene expression data, the algorithms need to scale to thousands of objects. We perform all runtime experiments on a laptop with an Intel Core i5-4300U CPU at 1.90 GHz (only using a single core), 8 GB RAM and running R 3.2.3 on Ubuntu 15.10. Figure 5 shows a summary of the runtimes for each algorithm on all datasets (random, pre anti-Robinson and real-world data). Obviously, identity and random are the fastest since they do not perform any seriation. The next fastest algorithms are based on hierarchical clustering with reordering. MDS, spectral seriation and TSP are in the middle followed by R2E, several QAP formulations and Dendser. The slowest are ARSA and the SPIN version. Note however, that the SPIN algorithms are implemented purely in R (an interpreted language) and thus are at a disadvantage against the others which are at least partially implemented in much faster C or FORTRAN.

To compare performance in terms of both, the quality of the seriation and speed, we use again all datasets and calculate average runtimes and the average gap to the best results found. Figure 6 shows all methods by average speed and gap. Efficient gap/speed combinations are marked and annotated with the method's name. The best results for AR events are produced by QAP formulations which take on average 20 ms. They outperform ARSA by producing similar quality but an order of magnitude faster. Spectral seriation and metric MDS and spectral seriation, which have a gap below 20%, are much faster with an average runtime around 5 ms. Dendrogram-based methods are fast, but produce inferior results with a gap greater than 40%. For path length, TSP produces the best results followed by hierarchical clustering with optimal leaf ordering (OLO) and various hierarchical clustering methods.

So far we have only presented results with very small datasets of around 100 objects, and we know that the worst case complexity of seriation is  $O(n!)$ . To investigate scalability to larger datasets we use again random data. We start with 100 objects and

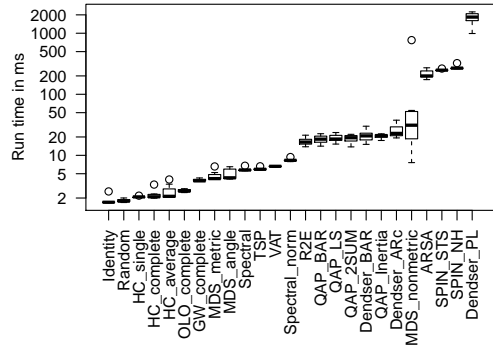


Figure 5: Comparison of runtime on all datasets.

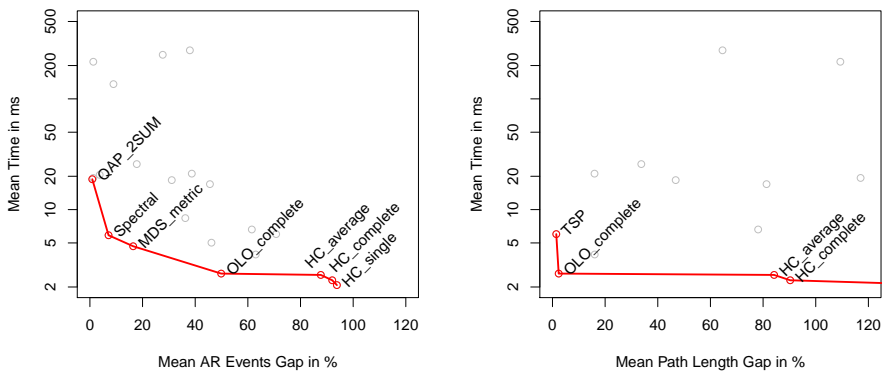


Figure 6: Comparison of method efficiency in terms of the speed/gap tradeoff.

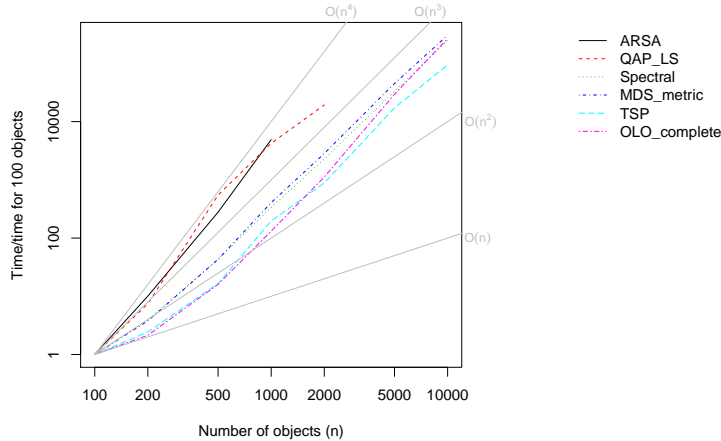


Figure 7: Scalability results for selected seriation methods.

475 then we double the number of object in each run. We only report here the results for  
a representative set of methods chosen from the efficient methods and close runners up  
identified above. We use a time limit of 5 minutes per run and stop once all methods run  
out of time. The results are shown in a log-log plot in Figure 7. To make methods using  
different programming languages comparable, we normalized runtime by the time it takes  
480 to seriate 100 objects. We also added grey lines for complexity of  $O(n)$ ,  $O(n^2)$ ,  $O(n^3)$   
and  $O(n^4)$  for reference. ARSA and the QAP solver for the linear seriation criterion are  
the most expensive with a complexity close to  $O(n^4)$ . These methods also run out of  
the time limit first at around 1000 and 2000 objects, respectively. Spectral seriation and  
metric MDS are very similar to each other with complexity lower than  $O(n^3)$ . This is not  
485 surprising since both are based on eigenvalue decomposition. Both can seriate around  
10,000 objects in under 5 minutes. Interestingly, the used TSP solver and hierarchical  
clustering with optimal leaf ordering have similar runtime complexity starting out with  
close to linear complexity for very small problems and then pick up and get close to MDS  
and spectral seriation with  $O(n^3)$ . These two methods are still the fastest and can also  
490 seriate up to 10,000 objects within the time limit.

It is interesting to note that methods which perform equally well in terms of the  
seriation criterion also have very similar complexity. In conclusion, we found that for  
practical applications spectral seriation and metric MDS provide a good tradeoff between  
seriation quality and runtime for gradient condition and rank/dissimilarity agreement-  
495 based criteria, while hierarchical clustering with optimal leaf ordering provides a good  
tradeoff for path length. Researchers and practitioners can conduct experiment with  
their own data using the R extension package seriation (Hahsler et al., 2016).

## 6. Conclusion

In this paper we provided a review and an experimental comparison of the most  
500 popular seriation criteria and heuristic methods used for seriation of one-mode two-way  
data. While criteria by construction fall into three groups based on gradient conditions,

rank/dissimilarity agreement and path length, the experimental study suggests that in addition a fourth group with the linear seriation criterion (including related banded anti-Robinson form criteria) and the gradient-based weighted Anti-Robinson events exists. The study also shows that the grouping sorts the criteria from representing only local structure all the way to a strong emphasis on global structure. Depending on which is more important for the application, criteria from a different groups can be used.

The comparison of popular seriation methods shows that the methods based on hierarchical clustering produce very similar results. All methods based on direct optimization of seriation criteria plus some other methods (metric MDS and SPIN) produce also very consistent seriation orders, while all other methods (including a pure TSP solver) create very different seriation results. For gradient-based seriation, QAP-based methods produce the best quality, while metric MDS and spectral seriation are very competitive and scale for larger datasets of up to 10,000 objects in under 5 minutes. For path length, hierarchical clustering with optimal leaf ordering performs very well and also scales to up to 10,000 objects in under 5 minutes.

Different seriation criteria and seriation methods highlight different structural aspects of the data and thus might be useful to explore in order to detect patterns which can be used to support decision making. This is easy to do with the open source software (Hahsler et al., 2016) used for all experiments in this paper.

## Acknowledgements

The author would like to thank the anonymous reviewers and the editor for their valuable comments and suggestions which lead to a substantial improvement of this paper. Special thanks also go to (in alphabetical order) Michael Brusco, Christian Buchta, Denise Earle, Kurt Hornik, Catherine Hurley, Hans-Friedrich Köhn, Fionn Murtagh, Franz Rendl, Günther Sawitzki and Stephanie Stahl, who contributed code to the R extension package seriation.

## References

- Alter, O., Brown, P. O., & Botstein, D. (2000). Singular value decomposition for genome-wide expression data processing and modeling. *Proceedings of the National Academy of Sciences (PNAS)*, 97.
- Applegate, D., Bixby, R., Chvátal, V., & Cook, W. (2006). *Concorde TSP Solver*. URL: <http://www.tsp.gatech.edu/concorde/>.
- Arabie, P., & Hubert, L. J. (1996). An overview of combinatorial data analysis. In P. Arabie, L. J. Hubert, & G. D. Soete (Eds.), *Clustering and Classification* (pp. 5–63). River Edge, NJ: World Scientific.
- Atkins, J. E., Boman, E. G., & Hendrickson, B. (1999). A spectral algorithm for seriation and the consecutive ones problem. *SIAM Journal on Computing*, 28, 297–310. doi:10.1137/S0097539795285771.
- Bar-Joseph, Z., Demaine, E. D., Gifford, D. K., & Jaakkola, T. (2001). Fast optimal leaf ordering for hierarchical clustering. *Bioinformatics*, 17, 22–29.
- Barnard, S. T., Pothen, A., & Simon, H. D. (1993). A spectral algorithm for envelope reduction of sparse matrices. In *Proceedings of the 1993 ACM/IEEE Conference on Supercomputing* (pp. 493–502). New York, NY, USA: ACM.
- Bezdek, J., & Hathaway, R. (2002). VAT: A tool for visual assessment of (cluster) tendency. In *Proceedings of the 2002 International Joint Conference on Neural Networks (IJCNN '02)* (pp. 2225–2230).
- Boutsinas, B. (2013). Machine-part cell formation using biclustering. *European Journal of Operational Research*, 230, 563–572.
- Brusco, M., Köhn, H. F., & Stahl, S. (2008). Heuristic implementation of dynamic programming for matrix permutation problems in combinatorial data analysis. *Psychometrika*, 73, 503–522.

- 550 Brusco, M., & Stahl, S. (2005). *Branch-and-Bound Applications in Combinatorial Data Analysis*. Springer-Verlag.
- Brusco, M. J. (2002). Integer programming methods for seriation and unidimensional scaling of proximity matrices: A review and some extensions. *Journal of Classification*, *19*, 45–67. doi:10.1007/s00357-001-0032-z.
- 555 Burkard, R. E., Cela, E., Pardalos, P. M., & Pitsoulis, L. S. (1999). The quadratic assignment problem. In P. Pardalos, & D.-Z. Du (Eds.), *Handbook of Combinatorial Optimization* (pp. 1713–1809). Springer-Verlag.
- Burkard, R. E., & Rendl, F. (1984). A thermodynamically motivated simulation procedure for combinatorial optimization problems. *European Journal of Operational Research*, *17*, 169–174.
- Caraux, G., & Pinloche, S. (2005). Permutmatrix: A graphical environment to arrange gene expression 560 profiles in optimal linear order. *Bioinformatics*, *21*, 1280–1281.
- Carroll, D., & Arabie, P. (1980). Multidimensional scaling. *Annual Reviews Psychology*, *31*, 607–649.
- Chen, C.-H. (2002). Generalized association plots: Information visualization via iteratively generated correlation matrices. *Statistica Sinica*, *12*, 7–29.
- Davey, B. A., & Priestley, H. A. (1990). *Introduction to lattices and order*. Cambridge: Cambridge 565 University Press.
- Deutsch, S. B., & Martin, J. J. (1971). An ordering algorithm for analysis of data arrays. *Operational Research*, *19*, 1350–1362.
- Ding, C., & He, X. (2004). Linearized cluster assignment via spectral ordering. In *Proceedings of the Twenty-first International Conference on Machine Learning (ICML '04)* (p. 30). ACM Press.
- 570 Earle, D., & Hurley, C. B. (2015). Advances in dendrogram seriation for application to visualization. *Journal of Computational and Graphical Statistics*, *24*, 1–25.
- Eisen, M. B., Spellman, P. T., Brownlager, P. O., & Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Science of the United States*, *95*, 14863–14868.
- 575 de Falguerolles, A., Friedrich, F., & Sawitzki, G. (1997). A tribute to J. Bertin’s graphical data analysis. In W. Bandilla, & F. Faulbaum (Eds.), *SoftStat '97: Advances in Statistical Software 6* (pp. 11–20). Lucius & Lucius.
- Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, *7*, 179–188.
- 580 Fogel, F., Aspremont, A. D., & Vojnovic, M. (2014). Serialrank: Spectral ranking using seriation. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, & K. Weinberger (Eds.), *Advances in Neural Information Processing Systems 27* (pp. 900–908). Curran Associates, Inc.
- Friendly, M. (2002). Corrgrams: Exploratory displays for correlation matrices. *The American Statistician*, *56*, 316–324.
- 585 Garfinkel, R. S. (1985). The traveling salesman problem: Motivation and modeling. In E. L. et al (Ed.), *The Traveling Salesman Problem* chapter 2. (pp. 17–36). John Wiley.
- George, A., & Pothen, A. (1997). An analysis of spectral envelope reduction via quadratic assignment problems. *SIAM Journal on Matrix Analysis and Applications*, *18*, 706–732.
- 590 Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. (1st ed.). Boston, MA, USA: Addison-Wesley Longman Publishing.
- Goulermas, J., Kostopoulos, A., & Mu, T. (2016). A new measure for analyzing and fusing sequences of objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *38*, 833–848. doi:10.1109/TPAMI.2015.2470671.
- Gruvaeus, G., & Wainer, H. (1972). Two additions to hierarchical cluster analysis. *British Journal of 595 Mathematical and Statistical Psychology*, *25*, 200–206.
- Gutin, G., & Punnen, A. P. (Eds.) (2002). *The Traveling Salesman Problem and Its Variations* volume 12 of *Combinatorial Optimization*. Dordrecht: Kluwer.
- Hahsler, M., Buchta, C., & Hornik, K. (2016). *Infrastructure for seriation*. URL: <http://CRAN.R-project.org/package=seriation> R package version 1.2-0.
- 600 Hahsler, M., & Hornik, K. (2016). *TSP: Traveling Salesperson Problem (TSP)*. URL: <http://CRAN.R-project.org/package=TSP> R package version 1.1-4.
- Hahsler, M., Hornik, K., & Buchta, C. (2008). Getting things in order: An introduction to the R package seriation. *Journal of Statistical Software*, *25*, 1–34. doi:10.18637/jss.v025.i03.
- 605 Hertzberg, M., Aspeborg, H., Schrader, J., A. Andersson, R., Blomqvist, K., Bhalerao, R., Uhlen, M., Teeri, T. T., Lundeberg, J., Sundberg, B., Nilsson, P., & Sandberg, G. (2001). A transcriptional roadmap to wood formation. *Proceedings of the National Academy of Sciences (PNAS)*, *98*, 14732–14737.

- Hodson, F. (1968). *The La Tene Cemetery at Münsingen-Rain*. Stämpfli, Bern.
- Holzinger, K. L., & Swineford, F. (1939). *A study in factor analysis: The stability of a bi-factor solution*.  
 610 Number 48 in Supplementary Educational Monograph. University of Chicago Press.
- Hubert, L., Arabie, P., & Meulman, J. (2001). *Combinatorial Data Analysis: Optimization by Dynamic Programming*. Society for Industrial Mathematics.
- Hubert, L., & Schultz, J. (1976). Quadratic assignment as a general data analysis strategy. *British Journal of Mathematical and Statistical Psychology*, *29*, 190–241.
- 615 Hubert, L. J. (1974). Some applications of graph theory and related nonmetric techniques to problems of approximate seriation: The case of symmetric proximity measures. *British Journal of Mathematical Statistics and Psychology*, *27*, 133–153.
- Kendall, D. (1971). Seriation from abundance matrices. In F. Hodson, D. Kendall, & P. Tautu (Eds.), *Mathematics in the Archaeological and Historical Sciences* (pp. 215–252). Edinburgh University Press.
- 620 Kendall, M. (1938). A new measure of rank correlation. *Biometrika*, *30*, 81–89.
- Laurent, M., & Seminaroti, M. (2015). The quadratic assignment problem is easy for Robinsonian matrices with Toeplitz structure. *Operations Research Letters*, *43*, 103–109.
- Lenstra, J. K. (1974). Clustering a data array and the traveling-salesman problem. *Operations Research*, *22*, 413–414.
- 625 Leskovec, J., & Krevl, A. (2014). SNAP Datasets: Stanford large network dataset collection. <http://snap.stanford.edu/data>.
- Lichman, M. (2013). UCI machine learning repository. URL: <http://archive.ics.uci.edu/ml>.
- Liiv, I. (2010). Seriation and matrix reordering methods: An historical overview. *Statistical Analysis and Data Mining*, *3*, 70–91.
- 630 Mair, P., & De Leeuw, J. (2015). Unidimensional scaling. In *Wiley StatsRef: Statistics Reference Online* (pp. 1–3). Wiley. doi:10.1002/9781118445112.stat06462.pub2.
- McCormick, W. T., Schweitzer, P. J., & White, T. W. (1972). Problem decomposition and data reorganization by a clustering technique. *Operations Research*, *20*, 993–1009.
- Mortenson, M. J., Doherty, N. F., & Robinson, S. (2015). Operational research from Taylorism to terabytes: A research agenda for the analytics age. *European Journal of Operational Research*, *241*, 583–595.
- 635 Paydar, M. M., & Saidi-Mehrabad, M. (2013). A hybrid genetic-variable neighborhood search algorithm for the cell formation problem based on grouping efficacy. *Computers & Operations Research*, *40*, 980–990.
- 640 Petrie, F. W. M. (1899). Sequences in prehistoric remains. *Journal of the Anthropological Institute*, *29*, 295–301.
- Robinson, W. S. (1951). A method for chronologically ordering archaeological deposits. *American Antiquity*, *16*, 293–301.
- Rogers, D. F., & Kulkarni, S. S. (2005). Optimal bivariate clustering and a genetic algorithm with an application in cellular manufacturing. *European Journal of Operational Research*, *160*, 423–440.
- 645 Sirovich, L. (2003). A pattern analysis of the second Rehnquist U.S. Supreme Court. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 7432–7437. doi:10.1073/pnas.1132164100.
- Soltysiak, A., & Jaskulski, P. (1998). Czekanowski’s diagram: A method of multidimensional clustering. In J. Barcelo, I. Briz, & A. Vila (Eds.), *Proceedings of the 26th Conference on Computer Applications and Quantitative Methods in Archaeology* (pp. 175–184). Barcelona, Spain: Archaeopress.
- 650 Thanh, L. T., Ferland, J. A., Elbenani, B., Thuc, N. D., & Nguyen, V. H. (2016). A computational study of hybrid approaches of metaheuristic algorithms for the cell formation problem. *Journal of the Operational Research Society*, *67*, 20–36.
- 655 Tien, Y.-J., Lee, Y.-S., Wu, H.-M., & Chen, C.-H. (2008). Methods for simultaneously identifying coherent local clusters with smooth global patterns in gene expression profiles. *BMC Bioinformatics*, *9*, 1–16.
- Tsafrir, D., Tsafrir, I., Ein-Dor, L., Zuk, O., Notterman, D. A., & Domany, E. (2005). Sorting points into neighborhoods (SPIN): Data analysis and visualization by ordering distance matrices. *Bioinformatics*, *21*, 2301–2308.
- 660 Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, *58*, 236–244.
- Wilkinson, E. (1971). Archaeological seriation and the travelling salesman problem. In F. Hodson, D. Kendall, & P. Tautu (Eds.), *Mathematics in the Archaeological and Historical Sciences* (pp. 276–283). Edinburgh University Press.
- 665 Wu, T.-H., Chung, S.-H., & Chang, C.-C. (2010). A water flow-like algorithm for manufacturing cell

formation problems. *European Journal of Operational Research*, 205, 346–360.

Yang, M.-S., & Yang, J.-H. (2008). Machine-part cell formation in group technology using a modified ART1 method. *European Journal of Operational Research*, 188, 140–152.